Reprinted without change of pagination from
IEEE Transactions of the Professional Group on Information Theory, Vol. IT-9, No. 3, July, 1963

5p

N64 13394 #
code none
Cr 53040

Technical Report No. 32-475

Optimal Search Procedures

Edward C. Posner 1 Jul. 1963 5p refs Repr. from

(NASA Contract NAS7-100)
(NASA CR-53040; JPL-TR-32-475)

This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration.

Reprint

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

July 1, 1963

Optimal Search Procedures*

EDWARD C. POSNER†, MEMBER, IEEE

13394

Summary—This paper sets up a restricted class of search procedures for a satellite lost in a region of the sky. The satellite must be found by a radar search. The procedures under consideration allow the use of a preliminary search, which may be done with a wider beam than is required for the final search. The purpose of the preliminary search is to obtain a ranking of the various portions of the sky, so that the final search can examine the more likely regions of the sky first. It is shown that a preliminary search can reduce the expected search time, with no matter how wide a beam it is carried out. It is also shown that the preliminary search with the narrowest possible beam is best. Author

SUPPOSE ONE is searching for a satellite or space probe with a radar beam. Suppose that the satellite is known to be in a certain portion of the sky composed of say A "cells." *A priori*, the satellite is as likely to be in any one cell as in any other; the radar beam must stay on the "correct" cell an amount of time t_1, say, in order to find the satellite and then track it if it is there. One must find the correct cell by this process. One method of search is to search each cell for the time t_1 just once until the object is found; it is the purpose of this paper to examine the consequences of other methods of search.

Consider search procedures of the following type. We allow ourselves to search regions of the sky of more than one cell by *widening the antenna beam*. Also, we need not at first search each cell for time t_1 (even if the beam isn't widened), but may make *preliminary searches* for less time and go back later to search for time t_1. We assume a radar system in which the averaged return from an individual cell is $\mu(> 0)$ when the satellite is present in the cell, 0 when it is absent. Additive white Gaussian noise of the same spectral density for different cells, and independent for different cells, is in the system; the noise has the effect of adding to the signal a Gaussian variable of variance σ^2 when an individual cell is observed for time t_0, say. The parameter μ/σ is assumed known; this is not an unreasonable assumption in view of the fact that one knows about how far away the satellite is, and also the approximate system temperature. The noise in the system is due mainly to receiver temperature. The "sky noise" is considered of secondary importance in this paper.

Under the above assumptions, when searching a region of k cells for time t, the signal decreases by a factor of $1/k$, whereas the noise variance changes by a factor of t_0/t. Thus, we may regard the signal as μ in the search with k cells clumped if the satellite is present in the region, and 0 if it is absent from the region, if we regard the noise variance as $k\sigma^2 t_0/t$. We first consider strategies of the following special type, called two-stage procedures.

Divide the A-celled portion of the sky in which the satellite is known to be into n *regions* of $m = A/n$ cells each. Then search each of these n regions for time t each, where t is to be determined. Rank the n regions in order of decreasing return, with the region giving the highest return first. Search each region cell by cell for time t_1, starting with the first region. The problem is to *choose n and t so as to minimize the expected search time.*

We first observe that we should take $n = A$. For the equivalent noise variance (the signal being regarded as constant), when searching a region of m cells for time t, is the same as the equivalent noise variance obtained in the system which searches each of the m cells of the region for time $\tau = t/m$ and then adds the returns. This follows from the fact that the variance of a sum of independent random variables is the sum of the variances. But having added the m returns, we would be discarding the *extra* information as to which of the m cells had the highest return. In the clumped procedure we would be searching the m cells in *random* order, whereas in the unclumped procedure we search all the cells in order of decreasing likelihood. Since the equivalent noise variances are the same in either case, the unclumped search must yield a smaller expected search time. A similar argument shows that it is better to ignore the regional boundaries, namely, to rank *all* the cells in order of decreasing return and search the cell with highest return first, the cell with second highest return second, etc., even if they occur in different regions. For once we know that it is better to "remember" the returns of the individual cells, a search which examines the more likely cells first yields a lower expected search time than any other procedure. Now consider a procedure which reverses the ranking which two cells i and j received according to their return; that is, suppose cell i had a higher return than cell j, but that we are given a procedure which examines j before it examines i. In the definition of expected search time as an expected value, the two procedures agree everywhere except in the terms involving cell i and cell j. Suppose cell i is the ith cell in order of decreasing return, and cell j the jth with $i < j$. The sum for the expected search time in the given case has terms it_1 pr (satellite is in cell j) + jt_1 pr (satellite is in cell i), it_1 being the time it takes to find the satellite if it is in fact in cell j under this procedure. The procedure to be proved optimum has instead the terms it_1 pr (satellite is in cell i) + jt_1 pr (satellite is in cell j). Since pr (satellite is in cell i) > pr (satellite is in cell j), because cell i had a higher return, the second sum is smaller than the first.

* Received July 2, 1962; revised manuscript received November 9, 1962. The research described here was conducted at the Jet Propulsion Laboratory, California Institute of Technology, under a program sponsored by the National Aeronautics and Space Administration, Contract No. NAS 7-100.

† Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Calif.

Thus, given any *two* cells, it is better to search them in order of decreasing return. Since *any* permutation of $1, 2, \cdots, n$ can be obtained by interchanging two indexes at a time, the search procedure which searches each cell in order of decreasing return is better than a procedure which searches the cells in any other order. This proves the assertion.

In this paper, we are ignoring the time it takes to move the radar beam from cell to cell; however, this time may be crucial since the cells are scattered in the portion of the sky by the ranking they have obtained from the preliminary search. If n is taken less than A, this difficulty may be avoided by using a spiral scan of the regions of A/n cells; the beam must then be moved only after the spiral scan of the region is completed. In fact, the preliminary search could perhaps be carried out with a lighter, wider beamed, finder antenna (this is analogous to the use of a finder telescope in optical astronomy). We shall see that a saving is possible over the method which searches each cell for time t_1 with no preliminary search, for *any* n, $2 \leq n \leq A$, and that one can get near the maximum savings (which we have proven occurs when $n = A$) with fairly small n.

Fixing n and t, we shall get a formula for the expected search time $E(t)$ and then minimize $E(t)$ with respect to t, keeping n fixed. Now $E(t)$ is a sum of three terms. The first is nt, the price we pay in the preliminary search process. The second is a term equal to the expected time of search of the region containing the correct cell. This term is $\sum_{i=1}^{m} it_1$ pr (cell i is correct, given that the region in question contains the correct cell)

$$= \sum_{i=1}^{m} it_1/m = t_1 \frac{(m+1)}{2}(m)\cdot\frac{1}{m} = t_1 \frac{(m+1)}{2}.$$

When $n = 1$, this gives $t_1((A+1)/2)$ as the expected search time without preliminary search. We thus must prove that the minimum of $E(t)$ in $t \geq 0$ for each n, $2 \leq n \leq A$ is less than $t_1((A+1)/2)$. The third and most interesting term contributing to the expected search time is the time wasted in searching those regions, whose returns have by chance come out higher than the return of the region containing the correct cell. This third term may be computed as follows. It equals the expected number of regions with a higher return than the region containing the correct cell times the time wasted in searching each of these incorrect regions; this wasted time is mt_1. The expected number of incorrect regions ranked before the correct one is the number of incorrect regions, $n-1$, times the probability that a given incorrect region is ranked before the correct one (even though the events in which different incorrect regions are ranked before the correct one are *not* independent). Thus, we must compute the probability that a given incorrect region is ranked before the correct one.

This reversal of position will occur if and only if the return from the incorrect region exceeds the return from the correct region; that is, if and only if a normal variate of mean 0, variance $m\sigma^2(t_0/t) = \sigma^2 t_0/\tau = \sigma_\tau^2$ say, exceeds an independent normal variate of mean μ, variance σ_τ^2, (if and only if the difference of these two variates is positive). Now the difference of two independent normal variates also has a normal density, whose mean is the difference of the two means and whose variance is the sum of the two variances. Thus we must compute the probability that a normal variate of mean $-\mu$, variance $2\sigma_\tau^2$, exceeds 0; we prefer to think of this as the probability that a unit normal variate z (mean 0, variance 1) exceeds $\mu/\sqrt{2}\sigma_\tau$, which we may write as pr $(z > \mu/\sqrt{2}\sigma_\tau)$. The term under discussion contributing to $E(t)$ is thus found to be

$$\frac{t_1}{m}(n-1)pr\left(z > \frac{\mu}{\sqrt{2}\sigma_\tau}\right).$$

We now have the required formula for $E(t)$,

$$E(t) = nt + t_1\left(\frac{m+1}{2}\right) + mt_1(n-1)pr\left(z > \frac{\mu}{\sqrt{2}\sigma_\tau}\right). \quad (1)$$

We shall now minimize $E(t)$. Differentiate $E(t)$ with respect to t; we obtain

$$E'(t) = n - mt_1(n-1)\phi\left(\frac{\mu}{\sqrt{2}\sigma_\tau}\right)\cdot\frac{d}{dt}\left(\frac{\mu}{\sqrt{2}\sigma_\tau}\right)$$

where $\phi(\mu)$ is the unit normal density function of μ. Using $\sigma_\tau = \sigma\sqrt{t_0/\tau}$, this becomes

$$E'(t) = n - mt_1(n-1)\phi\left(\frac{\mu}{\sqrt{2}\sigma_\tau}\right)\cdot\frac{\mu}{\sigma}\frac{1}{\sqrt{2mt_0}}\cdot\frac{1}{2\sqrt{t}}. \quad (2)$$

This formula exhibits $E'(t)$ as an increasing function of t in $t \geq 0$; thus $E(t)$ is a convex function of t in $t \geq 0$. Since $E(t) \to \infty$ as $t \to \infty$, we have shown that $E(t)$ has a unique minimum in $t \geq 0$ at, say $\bar{t}$.
But

$$E(0) = t_1\left(\frac{m+1}{2}\right) + mt_1(n-1)pr(z > 0)$$

$$= t_1\left(\frac{m+1}{2}\right) + mt_1\left(\frac{n-1}{2}\right) = t_1\left(\frac{A+1}{2}\right).$$

(This formula can also be proved by noting that the expected search time in the two-stage process as $t \to 0$ must approach the expected search time in the method which examines each cell with no preliminary search.) Also $E'(0) = -\infty$, so that $E(t)$ is less than $t_1((A+1)/2)$ near zero. This proves that $\bar{t} > 0$ and thus proves that a saving is possible for any n, $2 \leq n \leq A$.

We shall now find $\bar{t}$. Rewrite (2) as

$$E'(t) = n - \tfrac{1}{2}mt_1(n-1)\phi\left(\frac{\mu}{\sqrt{2}\sigma_\tau}\right)$$

$$\cdot\frac{\mu}{\sigma}\frac{1}{\sqrt{2mt_0}}\frac{1}{\frac{\mu}{\sigma}\sqrt{\frac{\tau}{2t_0}}}\frac{\mu}{\sigma}\frac{1}{\sqrt{2mt_0}} \quad (3)$$

then $\bar{t}$ satisfies the following equation where $\bar{\tau} = \bar{t}/m$:

$$n = \tfrac{1}{2} m t_1 (n-1) \phi\left(\frac{\mu}{\sqrt{2}\sigma_{\bar{\tau}}}\right)\left(\frac{\mu}{\sigma}\right)^2 \frac{1}{2mt_0} \frac{1}{\frac{\mu}{\sigma}\sqrt{\frac{2t_0}{\bar{\tau}}}},$$

or

$$\frac{\mu}{\sqrt{2}\sigma_{\bar{\tau}}} \phi\left(\frac{\mu}{\sqrt{2}\sigma_{\bar{\tau}}}\right) = \frac{\frac{4n}{n-1}}{\frac{t_1}{t_0}\left(\frac{\mu}{\sigma}\right)^2}. \tag{4}$$

Define $x = x(n, \mu/\sigma, t_1/t_0)$ by

$$\frac{1}{x}\phi(x) = \frac{\frac{4n}{n-1}}{\frac{t_1}{t_0}\left(\frac{\mu}{\sigma}\right)^2}; \tag{5}$$

then $x = \mu/\sqrt{2}\sigma_{\bar{\tau}}$, or

$$\bar{t} = \frac{2mt_0x^2}{\left(\frac{\mu}{\sigma}\right)^2}. \tag{6}$$

The expected saving S using this procedure instead of the method which searches each cell with no preliminary search is

$$S = t_1\left(\frac{A+1}{2}\right) - n\bar{t} - t_1\left(\frac{m+1}{2}\right) - mt_1(n-1)pr(z > x),$$

or

$$S = t_1\left(\frac{A+1}{2}\right) - \frac{2At_0x^2}{\left(\frac{\mu}{\sigma}\right)^2} - mt_1(n-1)(\tfrac{1}{2} - pr(0 < z < x)),$$

$$S = mt_1(n-1)pr(0 < z < x) - \frac{2At_0x^2}{\left(\frac{\mu}{\sigma}\right)^2}. \tag{7}$$

From this it can be proved that S increases with increasing n, $2 \leq n \leq A$. Since x clearly increases with n, we see that $n\bar{t}$, the total time spent in the preliminary search, also increases with n, being proportional to x^2. However, $\bar{t}$ decreases with n; we shall not prove this here. We also define the *fractional* saving

$$f = \frac{S}{t_1\left(\frac{A+1}{2}\right)};$$

thus,

$$f = \frac{2A}{A+1}\left[\left(\frac{n-1}{n}\right)pr(0 < z < x) - \frac{2x^2t_0}{\left(\frac{\mu}{\sigma}\right)^2 t_1}\right]. \tag{8}$$

Since the solution of (5) x does not depend upon A, we see that f hardly depends upon A.

An important point to bring out now is this; we can never find the satellite with certainty no matter how large t_1 is in comparison with t_0, that is, no matter how small the variance of the noise is in a search for time t_1. In practice one accepts a *confidence level* $\alpha(0 < \alpha < 1)$; that is, one chooses t_1 so large that the probability of saying the satellite is present in a cell when it is absent is $1 - \alpha$, or of saying that it is absent when it is present is $1 - \alpha$. The level α is chosen so small (in comparison with A) that the effect on the search of an error at any cell of either kind is negligible. One says that the satellite is *present* if the return exceeds $\mu/2$ and is *absent* if the return is less than $\mu/2$ in the time t_1; this is the maximum likelihood criterion. The probability of either kind of error is then $pr(z > (\mu/2)/\sigma\sqrt{t_0/t_1})$, where z has the unit normal density. For confidence level α, we must have

$$pr\left(z > \mu/2 \Big/ \sigma\sqrt{\frac{t_0}{t_1}}\right) = 1 - \alpha. \tag{9}$$

Define $z(\alpha)$ by

$$pr(z > z(\alpha)) = 1 - \alpha; \tag{10}$$

then $(\mu/2)/\sigma t_0/t_1 = z(\alpha)$, or

$$\frac{t_1}{t_0}\left(\frac{\mu}{\sigma}\right)^2 = 4z^2(\alpha). \tag{11}$$

Consider (8) as defining f as a function $f(n)$ of n. We see from (5) that since

$$\frac{1}{x}\phi(x) = \frac{\frac{n}{n-1}}{z^2(\alpha)}, \tag{12}$$

x depends only on n and α, but not on any other parameters. Eq. (8) becomes

$$f(n) = \frac{2A}{A+1}\left(\left(\frac{n-1}{n}\right)pr(0 < z < x) - \frac{x^2}{2z^2(\alpha)}\right), \tag{13}$$

where x is given by (12). Thus $f(n)$ depends soley on n, A and α, and not on any of the other parameters. When $\alpha = 0.999$, $z(\alpha) = 3.10$ and we have, from (11),

$$\frac{t_1}{t_0}\left(\frac{\mu}{\sigma}\right)^2 = 38.44;$$

This determines the signal-to-noise ratio required.

Fig. 1 graphs $f(n)$ vs n for $\alpha = 0.999$ and $A = 64$. We see that if n is not too small, $f(n)$ does not change much. For $n = 8$, we find $\bar{t}/t_1 = 0.45$; $8\bar{t}$, the total time spent in preliminary search, $= 3.6t_1$ and $f = 0.51$. For $n = 64$, we find $\bar{t}/t_1 = 0.0735$, $64\bar{t} = 4.7t_1$ and $f = 0.67$, the maximum possible saving with this method and these parameters and only $\frac{1}{3}$ more saving than the case $n = 8$. We see that with these cases the preliminary search corresponds to a "quick look," in that the total time spent in preliminary search amounts to only a few final cells' worth of time.

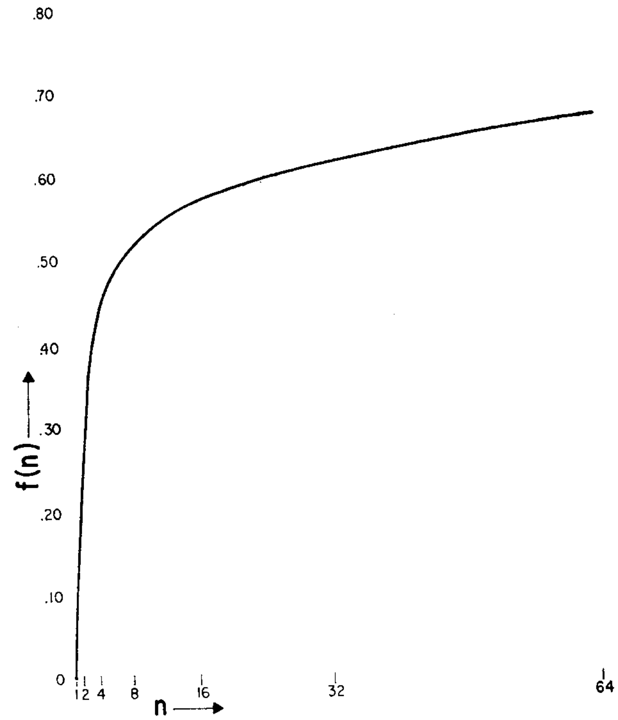

Fig. 1—$f(n)$ vs n.

We now discuss multistage, as opposed to two-stage, procedures. Here there are several kinds of aggregates of cells: cells, regions, regions of regions, etc. It is assumed that no "memory" is left from previous stages, but only the ordering is recorded. We claim the following: the optimal multistage procedure is the two-stage procedure with $n = A$. For proof, suppose we have a multistage procedure with $s > 2$ stages. We shall find a two-stage procedure with smaller expected search time. The last stage is by definition always a cell by cell search of all the A cells, in some order depending on the results of the previous stages, for time t_1 chosen to insure the given confidence level. We shall call the regions which occur in stage $s - i$ regions of type $s - i$. Consider the region of type $s - 2$ containing the correct cell. We know from the argument given at the beginning of this paper that it is better to search the region of type $s - 2$ by a preliminary search of each cell rather than by a preliminary search of regions of more than one cell obtained by clumping. That is, we may assume that the given procedure has regions of type $s - 1$ consisting of one cell each. As before, it is still better to ignore the boundaries of the regions of type $s - 2$ in conducting the final cell by cell search for time t_1 per cell; rather, one should rank the returns of *all* the cells at stage $s - 1$, regardless of which region of type $s - 2$ they are in, and then search each cell for time t_1 at the final stage. This means that the rankings obtained from stages $s - 2$ and earlier are not used at all in the final search. This proves that there is a two-stage procedure with smaller expected search time.

One can also consider the case of n-multiple finder antennas. Here the preliminary search is carried out on all the n regions simultaneously so that even more savings can be realized. This system may be more easily implementable than the system which uses the original antenna as a finder antenna. The mathematical treatment is similar, the only difference comes from an initial term of t in (1) instead of nt. Define x_2 by the analogue of (5):

$$\frac{1}{x_2}\,\phi(x_2) = \frac{\dfrac{4n}{n-1}}{\dfrac{t_1}{t_0}\left(\dfrac{\mu}{\sigma}\right)^2}.$$

Calling the optimum preliminary search time $\bar{\bar{t}}$, one then has

$$\bar{\bar{t}} = \frac{2mt_0}{\left(\dfrac{\mu}{\sigma}\right)^2}.$$

The details are omitted.

This type of search problem can arise in other contexts. For example, one may be searching for a narrow-band radar return lost in part of a wider, noisy, Doppler band. The same methods and results would apply. See [1].

It is to be mentioned that most types of search problems consider the minimization of the number of *steps* required to find a lost or hidden object. Here we have minimized the expected *time* of search. The question arises as to what the *optimum* strategy for minimizing the expected search time would be in the situation described in this paper. Here there is no restriction that the final search be a cell-by-cell search for the same time t_1 per cell, or even that the search be broken down into well defined stages. Furthermore, all the information gathered on the previous searches is to be used, not only rankings. In a paper now in preparation by the author and H. C. Rumsey, Jr., it will be shown that the optimal strategy for minimizing the expected search time is as follows, even without the assumption that each cell is *a priori* equally likely. Search the most likely cell until it is no longer the most likely, then search the cell which has become the most likely. A formula is found for the minimum; it looks something like the entropy of the *a priori* distribution, but is not proportional to the entropy. This will give another measure of the uncertainty of a finite probability distribution, the relevant definition for the search problem. These ideas will be examined further in the above-mentioned paper.

Bibliography

[1] A. A. Galvin, "A sequential detection system for the processing of radar returns," Proc. IRE, vol. 49, pp. 1417–1423; September, 1961.

Reprinted from the IEEE TRANSACTIONS OF THE PROFESSIONAL TECHNICAL GROUP ON *INFORMATION THEORY*
Volume IT-9, Number 3, July 1963
PRINTED IN THE U.S.A.